Article Watch, July 2013

Clive A. Slaughter

GRU-UGA Medical Partnership, Athens GA

This column highlights recently published articles that are of interest to the readership of this publication. We encourage ABRF members to forward information on articles they feel are important and useful to Clive Slaughter, Georgia Regents University/University of Georgia Medical Partnership, 1425 Prince Ave., Athens, GA 30606. Phone: 706-713-2216; Fax: 706-713-2221; E-mail: cslaught@uga.edu; or to any member of the editorial board. Article summaries reflect the reviewer's opinions and not necessarily those of the association.

DNA SEQUENCING AND CHARACTERIZATION

Ramasamy A, Trabzuni D, Gibbs J R, Dillman A, Hernandez D G, Arepalli S, Walker R, Smith C, Ilori G P, Shabalin A A, Li Y, Singleton A B, Cookson M R, for NABEC, Hardy J, for UKBEC, Ryten M, Weale M E. Resolving the polymorphism-in-probe problem is critical for correct interpretation of expression QTL studies. *Nucleic Acids Research* 41;2013:e88.

Studies of expression quantitative trait loci (eQTL) seek associations between single nucleotide polymorphisms (SNPs) and gene-expression levels. Many of these eQTL studies have used expression microarrays to quantify gene-expression levels, but a problem arises when the probes on the microarrays, which are designed to match a single reference sequence, contain SNPs within the probe sequence. Sequences that differ from the reference as a result of polymorphism give weaker signals when tested on the microarray and therefore, produce an apparent association between expression level and genotype. Moreover, polymorphic loci in linkage disequilibrium with such SNPs also display an association. The treatment of this polymorphism-in-probe problem has been uneven. The present study documents the impact of the problem and concludes that polymorphisms-in-probe massively inflate the number of cis-eQTL signals and that effect sizes of false eQTL signals are large compared with true signals. The authors suggest a way to resolve the problem definitively. With the use of the latest genome and exome reference data, nearly all polymorphisms (i.e., common variants) that reside within probe sequences can be identified and excluded from consideration. The authors do so for two of the most popularly used microarrays: the Illumina Human HT-12 array and the Affymetrix Human Exon 1.0 ST array.

Zong C, Lu S, Chapman A R, Xie X S. Genome-wide detection of single-nucleotide and copy-number variations of a single human cell. *Science* 338;2012:1622–1626.

Single-cell sequencing promises to allow investigation of genomic stability and cancer development, but the realization of this promise has been impeded by bias and error that are introduced when the picogram amounts of DNA in a single cell are amplified to the microgram amounts needed for sequencing. A major advance in the fidelity of whole genome amplification was achieved when the earliest amplification methods based on PCR were replaced by multiple displacement amplification (MDA), a technique that incorporates random priming and strand displacement by φ29 polymerase. Although much improved, the level of bias with MDA still precludes accurate genotyping from single cells, an application that depends on a high probability that both alleles will be amplified if heterozygotes are to be reliably recognized. Zong et al. have developed a new process for whole-genome amplification that promises further improvements in amplification fidelity. The process is called multiple annealing and looping-based amplification cycles (MALBAC). It uses primers that incorporate an 8-nt segment for random priming plus a common 27-nt sequence. Multiple short cycles of amplification, catalyzed by *Bst* polymerase and primed by the random segment, are interspersed by melting and reannealing to promote more even genome coverage. Complete amplicons contain the 27-nt tag at one end and its complementary sequence at the other end. When melted off the template, they therefore form loops that resist further hybridization and amplification. These full amplicons are subsequently amplified by PCR. The efficiency of detecting both alleles in heterozygotes by MALBAC is 71% compared with only 10% by MDA. The results of the work demonstrate how amplification bias can be ameliorated by controlling the priming phase of the process.

MACROMOLECULAR SYNTHESIS AND SYNTHETIC BIOLOGY

Mali P, Yang L, Esvelt K M, Aach J, Guell M, DiCarlo J E, Norville J E, Church G M. RNA-guided human genome engineering via Cas9. *Science* 339;2013:823–826.

doi: 10.7171/jbt.13-2402-006

Cong L, Ran F A, Cox D, Lin S, Barretto R, Habib N, Hsu P D, Wu X, Jiang W, Marraffini L A, Zhang F. Multiplex Genome engineering using CRISPR/Cas systems. *Science* 339;2013:819–823.

Cho S W, Kim S, Kim J M, Kim J-S. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nature Biotechnology* 31;2013:230–232.

Hwang W Y, Fu Y, Reyon D, Maeder M L, Tsai S Q, Sander J D, Peterson R T, Yeh J R, Joung J K. Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nature Biotechnology* 31;2013:227–229.

Jiang W, Bikard D, Cox D, Zhang F, Marraffini L A. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nature Biotechnology* 31;2013:233–239.

Targeted editing of specific genes is a technology of considerable promise, not only for experimental work but also for clinical application. Until very recently, the use of zinc-finger nucleases has dominated this field, but in 2009, the transcription activator-like effector proteins were developed into effector nucleases (TALENs) that are now competing with zinc-finger nucleases. (See also *Functional Genomics and Proteomics* below.) The present papers mark the appearance of a further kind of customizable molecular scissors, guided by RNA to their target genes. The new system is derived from a component of bacterial and archaeal immunity that is based on the clustered, regularly interspaced, short palindromic repeats (CRISPR) interference system. In this system, the CRISPR-associated (Cas) endonuclease cleaves foreign nucleic acids, such as bacteriophage genomes, directed to its target sequence by two small RNA molecules, to enable it to cleave the foreign DNA. Unlike genome editing with zinc-finger nucleases and TALENs, customizable targeting of Cas endonuclease requires only the design of new RNA guides, not a new pair of enzymes. Mali et al. and Cong et al. report mutations introduced at the cleavage site by nonhomologous end-joining and sequence introduction from a donor construct by homologous recombination. Cho et al. derive mutant human cell lines by limiting dilution following the use of Cas endonuclease without the need for antibiotic selection. Hwang et al. apply the technology to mutation of genes in zebrafish. Jiang et al. investigate the target specificity of the system. The development of RNA-targeted editing is widely believed to represent a breakthrough in efficiency and convenience for genome engineering.

SMALL MOLECULE ANALYSIS AND METABOLOMICS

Link H, Kochanowski K, Sauer U. Systematic identification of allosteric protein-metabolite interactions that control enzyme activity in vivo. *Nature Biotechnology* 31;2013:357–361.

A rapid way for organisms to respond to environmental change is to alter the activity of an enzyme or enzymes in response to changing pool sizes of metabolites with which the enzymes allosterically interact. Physiologically important allosteric interactions are usually inferred from computational modeling rather than from experiment or are simply assumed a priori. The present paper, however, provides an experimental approach to ascertaining relevant allosteric interactions. Culture conditions of *Escherichia coli* are switched every 30 s between medium containing pyruvate and medium containing ^{13}C-labeled glucose or fructose. This causes reversals between glycolysis and gluconeogenesis. Changes in metabolic flux are monitored by quantification of tracer labeling. Fifty-three metabolites are also quantified by mass spectrometry during the 30-s switches to ascertain the role of allosteric metabolite binding in the flux reversals. This perturbation approach here results in the identification of new allosteric effectors that determine the reversals of metabolic flux in the system, including one in which pyruvate activates fructose-1,6-bisphosphatase. The general approach can be used to characterize allosteric regulation but may also prove useful in identifying the targets of small-molecule inhibitors.

Tulipani S, Llorach R, Urpi-Sarda M, Andres-Lacueva C. Comparative analysis of sample preparation methods to handle the mplexity of the blood fluid metabolome: when less is more. *Analytical Chemistry* 85;2012:341–348.

The accuracy of metabolite fingerprinting of blood fluid samples by liquid chromatography-mass spectrometry (LC-MS) depends on removing substances that may interfere with the analysis by causing ion suppression or enhancement, by degrading the accuracy of mass measurements, or by binding to the analytes of interest. Blood proteins are known to cause such interferences and are removed routinely by organic solvent precipitation or by heating prior to LC-MS analysis. Phospholipids, however, also cause interference, but their removal prior to LC-MS analysis is not normally considered. As the organic solvent treatments that remove proteins do not also quantitatively extract phospholipids, the present paper describes the effect on detection of nonlipid metabolites of removing phospholipids by solid-phase extraction. The procedure is found to improve detection of most nonlipid, low molecular-weight metabolites. As blood-lipid content is highly variable with time, individual, and gender, this extra step is likely to improve overall variability in measurement of nonlipid metabolites.

CARBOHYDRATES AND GLYCOCONJUGATES

Kolarich D, Rapp E, Struwe W B, Haslam S M, Zaia J, McBride R, Agravat S, Campbell M P, Kato M, Ranzinger R, Kettner C, York W S. The Minimum Information Required for a Glycomics Experiment (MIRAGE) project: improving the standards for reporting mass-spectrometry-based glycoanalytic data. *Molecular & Cellular Proteomics* 12;2013:991–995.

Wells L, Hart G W. Glycomics: building upon proteomics to advance glycosciences. *Molecular & Cellular Proteomics* 12;2013:833–835.

The Minimum Information Required for a Glycomics Experiment (MIRAGE) project was initiated at a 2009 workshop on analytical and bioinformatic glycomics, organized by the Consortium for Functional Glycomics, an international group of glycoscientists. Wells and Hart, in their introduction to a series of review articles on glycomic methods in *Molecular & Cellular Proteomics*, further provide a summary of the journal's requirments for publishing experiments on glycoconjugate analysis.

Ni W, Bones J, Karger B L. In-depth characterization of N-linked oligosaccharides using fluoride-mediated negative ion microfluidic chip LC-MS. *Analytical Chemistry* 85;2013:3127–3135.

Chromatographic separation of N-glycans on porous, graphitized carbon is here combined with addition of ammonium fluoride to the chromatographic mobile phase to improve chromatographic and on-line mass spectrometric performance in structural analysis of N-linked sugars. The presence of fluoride helps recovery of sialylated oligosaccharides. Fluoride also assists the formation of negative ions by first, hydrogen bonding to C3 of N-acetyl glucosamine at the reducing end and then, abstracting a proton with neutral loss of HF. The result is a preponderance of $[M-H]^-$ ions. Collision-induced fragmentation of these ions generates C-type glycosidic fragments, structurally informative A-type cross-ring fragments, and D- and E-type fragments that in combination, are helpful in characterizing linkage and positional isomers.

MASS SPECTROMETRY

Inutan E D, Trimpin S. Matrix assisted ionization vacuum (MAIV), a new ionization method for biological materials analysis using mass spectrometry. *Molecular & Cellular Proteomics* 12;2013:792–796.

Trimpin S, Inutan E D. New ionization method for analysis on atmospheric pressure ionization mass spectrometers requiring only vacuum and matrix assistance. *Analytical Chemistry* 85;2013:2005–2009.

A new, soft ionization method, complementary to electrospray ionization (ESI) and matrix-assisted laser desorption/ionization (MALDI), is described and applied in these two papers. In the new method, called matrix-assisted ionization vacuum (MAIV), substances are transported directly from a solid support and ionized with the assistance of a matrix compound, 3-nitrobenzonitrile (3-NBN), which is known to sublime and also to triboluminesce (produce light when crystals are sheared or crushed). These properties may be important in the production of analyte ions by MAIV. Ionization occurs without application of high voltages, a laser beam, or heat. It merely requires reduced pressure from the vacuum of the mass spectrometer. The process can be initiated in an intermediate vacuum MALDI source with the laser turned off or in an electrospray source that is closed to atmospheric pressure. These conditions may be readily achieved by minor adaptations to existing, commercial ion sources. The spectra produced by MAIV are similar to ESI spectra, notably in the production of multiply charged ions of proteins. Sample preparation is similar to that for MALDI, except that a sample plate with many 3-NBN-treated samples cannot be introduced into the ion source on a single target, as they would all produce ions simultaneously. The new method is demonstrated in analyses of drugs, peptides, and proteins at the low femtomole level. It is suitable for analysis of dried-blood samples, filter paper strips, and tissue sections. The robustness and simplicity of instrument design requirements and the wide variety of substances amenable to analysis make the method potentially useful in clinical settings and in field-portable instruments.

PROTEINS—PURIFICATION AND CHARACTERIZATION

Zhao H, Ghirlando R, Piszczek G, Curth U, Brautigam C A, Schuck P. Recorded scan times can limit the accuracy of sedimentation coefficients in analytical ultracentrifugation. *Analytical Biochemistry* 437;2013:104–108.

Zhao et al. have detected systematic inaccuracies in the elapsed time recorded in data files from the analytical ultracentrifuge when using the manufacturer's latest data-acquisition software. The source of the error is unknown. It results in overestimates of sedimentation coefficients of up to 10% and resultant errors in hydrodynamic shape and molar mass estimates. Investigators are advised to re-examine recently published sedimentation coefficients. The problem can be recognized by comparing the elapsed scan times recorded in data-acquisition files with file time-stamps recorded by the instrument's operating system. A routine in SEDFIT has been written to alert users to discrepancies and to correct scan times.

Hagn F, Etzkorn M, Raschle T, Wagner G. Optimized phospholipid bilayer nanodiscs facilitate high-resolution structure determination of membrane proteins. *Journal of the American Chemical Society* 135;2013:1919–1925.

Phospholipid nanodiscs have been introduced recently as a new membrane mimetic environment for NMR analysis of integral membrane proteins in phospholipid bilayers. These nanodiscs are formed by wrapping apolipoprotein A-1 (Apo A-1) around small patches of a phospholipid bilayer, creating small, membrane-like particles of defined size. These structures avoid the use of detergents that may lower membrane–protein stability and abolish protein function. Versions of Apo A-1, called membrane scaffold proteins, have been engineered previously for use in this application but still produce inconveniently large nanodiscs. The present paper describes new, truncated versions of Apo A-1 that form smaller nanodiscs that permit improvements in solution-phase NMR spectral quality. It is hoped that these smaller nanodiscs will facilitate de novo structure determination of small- to medium-sized membrane proteins and enable studies of the interaction between membrane proteins and their soluble protein partners.

Rispens T, Heer P O, Derksen N I, Wolbink G, van Schouwenburg P A, Kruithof S, Aalberse R C. Nanomolar to sub-picomolar affinity measurements of antibody-antigen interactions and protein multimerizations: fluorescence-assisted high-performance liquid chromatography. *Analytical Biochemistry* 437;2013:118–122.

Binding affinities that are very high are difficult to measure by surface plasmon resonance, as dissociation times are very long. Rispens et al. here describe a method by which a dissociation constant for the interaction of an antibody and a small molecule ligand below 1 pM is measured successfully. The method uses HPLC with fluorescence detection. One binding partner is fluorescently labeled and incubated with varying concentrations of the other partner. Bound and free forms of the labeled protein are then separated by chromatography, and their peak heights provide estimates for the amounts bound. A dissociation constant may then be calculated. This approach demands that the separation time is short compared with the dissociation rate, but for interactions with dissociation constants in the nanomolar and picomolar ranges, this condition is usually met.

PROTEOMICS

Picotti P, Clement-Ziza M, Lam H, Campbell D S, Schmidt A, Deutsch E W, Rost H, Sun Z, Rinner O, Reiter L, Shen Q, Michaelson J J, Frei A, Alberti S, Kusebauch U, Wollscheid B, Moritz R L, Beyer A, Aebersold R. A complete mass-spectrometric map of the yeast proteome applied to quantitative trait analysis. *Nature* 494;2013:266–270.

Past attempts to assemble complete reference maps to cover an entire eukaryotic proteome have relied on the development of immunoassays or on shotgun proteomic experiments to compile spectral libraries. Neither method has succeeded in covering more than approximately two-thirds of the yeast proteome. Here, an almost complete reference map covering 97% of the proteome of *Saccharomyces cerevisiae* is assembled by combining high-throughput peptide synthesis with mass spectrometry (MS). Up to eight peptide sequences with favorable mass-spectrometric characteristics (proteotypic peptides) were selected from each of the 6607-predicted yeast-protein sequences. Two hundred proteins remained refractory to the selection criteria, as they lack suitable trypsin cleavage sites. The resulting library of ~28,000 peptides was synthesized on a small scale by the SPOT synthesis technique. Two reference spectral libraries were acquired using this set of peptides—one to support discovery proteomics using a linear ion-trap mass spectrometer and the other to support directed quantification by single-reaction monitoring on a QTRAP mass spectrometer. The final spectral libraries contained spectra for 100,815 and 28,216 peptides, respectively. Approximately 1630 peptides could not be detected by MS, probably as they were very hydrophobic or very hydrophilic and were therefore unsuitable for introduction via reverse-phase liquid chromatography. The proteome map generated in this study along with the tools to navigate it are publically available as a resource for proteomic analyses in yeast.

Rhee H-W, Zou P, Udeshi N D, Martell J D, Mootha V K, Carr S A, Ting A Y. Proteomic mapping of mitochondria in living cells via spatially restricted enzymatic tagging. *Science* 339;2013:1328–1331.

To assign subcellular localizations to proteins, the present study uses a peroxidase enzyme that can be genetically targeted to specific subcellular compartments and biotinylate nearby proteins for subsequent purification and identification by mass spectrometry. The enzyme, an ascorbate peroxidase, has a small labeling radius (<20 nm) in intact cells, providing specificity of labeling within subcellular regions. The methodology is used to investigate the localization of human mitochondrial proteins and results in identification of 495 proteins in the mitochondrial matrix, including 31 not previously linked to mitochondria. It distinguishes integral membrane proteins facing the intermembrane space from those facing the matrix. The results support reassignment to the mitochondrial matrix of several proteins previously thought to be localized to the outer membrane. These reassignments are confirmed by electron microscopy. Labeling is initiated by addition of biotin-

phenol and hydrogen peroxide to live cells. Labeling is sufficiently brief to support a temporal resolution of 1 min.

DRUG SCREENING

Gut P, Baeza-Raja B, Andersson O, Hasenkamp L, Hsiao J, Hesselson D, Akassoglou K, Verdin E, Hirschey M D, Stainier D Y R. Whole-organism screening for gluconeogenesis identifies activators of fasting metabolism. *Nature Chemical Biology* 9;2013:97–104.

High-throughput drug screening in rodents is impractical, while screening with cultured cells cannot target processes that integrate information across different cell types and tissues. Gut et al. demonstrate the use of zebrafish larvae to overcome these problems. They conduct a high-throughput screen for small molecules that affects energy metabolism, making use of the evolutionarily conserved architecture of energy homeostasis among metazoans. Zebrafish larvae derive carbohydrates from their yolk sac, but when this source is depleted, they switch to a "fasting" mode of metabolism in which gluconeogenesis maintains glucose levels. At this stage, the larvae have already developed functional organs yet are transparent and may be cultured in 96-well plates. During this fed-to-fasting transition, expression of the prototypical fasting-response protein, phosphoenolpyruvate carboxykinase (PEPCK), is induced. Gut et al. use transgenic bioluminescence and fluorescence reporter zebrafish to monitor gluconeogenesis. They validate the model for small-molecule screening by identifying U.S. Food and Drug Administration-approved drugs, with conserved actions on gluconeogenesis in zebrafish and humans, and by characterizing compounds with previously unknown actions on mammalian energy homeostasis. Their results suggest that the zebrafish-screening system is appropriate for identifying new drug targets for modulation of human metabolic diseases, such as diabetes and obesity.

FUNCTIONAL GENOMICS AND PROTEOMICS

Maeder M L, Linder S J, Reyon D, Angstman J F, Fu Y, Sander J D, Joung J K. Robust, synergistic regulation of human gene expression using TALE activators. *Nature Methods* 10;2013:243–245.

Perez-Pinera P, Ousterout D G, Brunger J M, Farin A M, Glass K A, Guilak F, Crawford G E, Hartemink A J, Gersbach C A. Synergistic and tunable human gene activation by combinations of synthetic transcription factors. *Nature Methods* 10;2013:239–242.

Transcription activator-like effectors (TALEs) were discovered as bacterial proteins that are injected into plant cells during infection by a bacterial pathogen. TALE proteins consist of an N-terminal translocation domain, a central repeat domain (there are usually 15.5–19.5 single repeats within this domain, each one consisting of ~34 highly conserved amino acid residues) that mediates binding to host DNA, and a C-terminal transcription activation domain. The amino acid sequence of the repeat determines its DNA-binding specificity: a particular pair of highly variable residues within the otherwise highly conserved repeat sequence determines its nucleotide target sequence. TALEs can be made that can bind specifically to any DNA and by linking the TALE DNA-binding domain to different functional protein domains, they can be made into highly specific transcriptional activators, repressors, or nucleases (the latter are called TALENs). However, the induction rates achieved with activator TALEs have hitherto been rather low. The present two papers show that synergistic combinations of different TALEs acting on the same target gene can often achieve much better induction. With the use of appropriate combinations of such mixtures, the transcription rate of target genes can be finely tuned, making these transcriptional switches effective over a wide dynamic range.

Zhu J, Larman H B, Gao G, Somwar R, Zhang Z, Laserson U, Ciccia A, Pavlova N, Church G, Zhang W, Kesari S, Elledge S J. Protein interaction discovery using parallel analysis of translated ORFs (PLATO). *Nature Biotechnology* 31;2013:331–334.

Zhu et al. describe a workflow for discovery of the proteins with which particular biomolecules of interest interact. The workflow involves ribosome display coupled with deep sequencing. A library of over 15,000 concentration-normalized human ORFs is used to prepare a library of mRNA molecules that lack stop codons. After translation, these mRNAs remain tethered to the proteins they encode. From this ribosome display library, an immobilized bait molecule then selects the proteins with which it interacts, and these are identified by deep sequencing of their attached mRNAs. This workflow is tested by identifying proteins that interact with LYN kinase, which contains common structural components of the Src family, with autoantibodies from patients with autoimmune disease, and with the drug molecules gefitinib and desatinib. The limitations of the methodology include incompleteness of the present library and lack of protein post-translational modifications. On the other hand, protein size and composition have minimal effects on display efficiency (in contrast to the problems encountered in expressing proteins for inclusion in microarrays). Also, cost and instrumentation requirements are low, and given steady improvements in the cost of DNA sequencing, the

methodology is becoming increasingly suitable for testing large numbers of samples.

CELL BIOLOGY AND TISSUE ENGINEERING

Bugaj L J, Choksi A T, Mesuda C K, Kane R S, Schaffer D V. Optogenetic protein clustering and signaling activation in mammalian cells. *Nature Methods* 10;2013:249–252.

Signaling processes commonly achieve the spatial and temporal control demanded of them by incorporating into one or more of the activation steps a molecular clustering or complex-assembly event that accelerates the response by increasing the local concentration of participating proteins. Bugaj et al. here present an experimental system that allows modular and tunable control of oligomerization in response to light and use it to control a signaling cascade in cultured cells. The optogenic response is based on the protein crypochrome 2 (Cry2) from *Arabidopsis*, which aggregates to form "photobodies" in plant cells in response to blue light and which disperses again when illumination ceases. Cry2 is here fused to lipoprotein receptor-related protein 6c, a coreceptor in the Wnt/β-catenin signaling pathway that forms clusters in response to signaling by the natural ligand, Wnt. The signal relieves β-catenin inhibition and promotes β-catenin nuclear translocation and transcriptional activity. Illumination is shown to achieve a transcriptional response even higher than that elicited by Wnt. The effect is demonstrated further by photoactivating the Rho GTPase, RhoA, to produce cytoskeletal contraction. It is hoped that this methodology will prove useful for optogenetic control of diverse signaling pathways in vivo.

Gomez D, Shankman L S, Nguyen A T, Owens G K. Detection of histone modifications at specific gene loci in single cells in histological sections. *Nature Methods* 10; 2013:171–177.

Ascertaining the contribution to cellular differentiation made by specific epigenetic marks at particular genes has hitherto been achieved by chromatin immunoprecipitation, a technique that can be deployed only on whole populations of cells. The technique is unsuitable for following specific epigenetic marks in individual cells embedded in tissues containing many cell types. Gomez et al. here describe methodology for visualizing a particular histone modification specific to smooth muscle cells (SMCs), dimethylation of lysine 4 in histone H3 (H3K4me2), at the myosin heavy-chain 11 (*MYH11*) gene locus, with single-cell resolution. The method relies on a variant of the proximity ligation assay, a technique used previously for detecting protein–protein interaction. A biotinylated DNA probe targets the gene of interest, *MYH11*, by in situ hybridization and is then bound by antibiotin antibodies. A different antibody probe targets the chromatin modification, H3K4me2. The two antibodies bear DNA aptamers that when immobilized in proximity, ligate to a linker. The ligation product is then detected by rolling circle amplification. The method is shown to detect the mark in the SMCs of blood vessels in tissues, including formaldehyde-fixed, paraffin-embedded tissue sections collected for clinical purposes. The mark is shown to persist in SMCs that have migrated to atherosclerotic lesions and have ceased to express SMC marker genes. The methodology is anticipated to find application in lineage tracing of cells in tissues and in studies to investigate the regulation of phenotypic transitions in diseases, such as atherosclerosis and cancer.

Sharei A, Zoldan J, Adamo A, Sim W Y, Cho N, Jackson E, Mao S, Schneider S, Han M-J, Lytton-Jean A, Basto P A, Jhunjhunwala S, Lee J, Heller D A, Kang J W, Hartoularos G C, Kim K-S, Anderson D G, Langer R, Jensen K F. A vector-free microfluidic platform for intracellular delivery. *Proceedings of the National Academy of Sciences USA* 110;2013:2082–2087.

There exists a great variety of ways to introduce macromolecules into cells experimentally. Here is a new method that entails mechanical deformation of cells as they pass through a constriction in a microfluidic channel that is 30–80% smaller than the cell diameter. Presumably, the controlled compression and sheer forces cause the transient appearance of holes through which molecules can diffuse from the surrounding medium. Electroporation and microinjection also rely on membrane disruption, but the present method is more successful at delivering proteins and less cytotoxic than electroporation and higher in throughput than microinjection. It does not rely on the endocytic pathway, as do many delivery methods, with the associated risk of trapping in endosomal or lysosomal vesicles; nor does it depend on chemical modification of cargo molecules to allow them to penetrate membranes. The new method proves successful in delivering carbon nanotubes, proteins, and small interfering RNA and is used on 11 cell types, including cells traditionally difficult to transfect. It is hoped to find application in diverse basic and clinical science applications.

IMAGING

Chung K, Wallace J, Kim S-Y, Kalyanasundaram S, Andalman A S, Davidson T J, Mirzabekov J J, Zalocusky K A, Mattis J, Denisin A K, Pak S, Bernstein H, Ramakrishnan C, Grosenick L, Gradinaru V, Deisseroth K. Structural and molecular interrogation of intact biological systems. *Nature* 497;2013:332–337.

Creating a three-dimensional structural image of the brain usually depends on reassembling images of serial thin sections, a cumbersome procedure that is susceptible to

error. Tissue lipids are primarily responsible for scattering light and thus, make the tissue optically opaque. Lipids are also a major barrier to the penetration of macromolecular probes. Yet, removal of lipids to render tissues transparent and permeable causes major disruption of tissue architecture. Chung et al. here describe a new methodology called CLARITY that overcomes this problem. Intact mouse brains are infused with formaldehyde and with acrylamide and bis-acrylamide monomers. The formaldehyde stabilizes the tissue by forming cross-links between tissue components and links acrylamide monomers to tissue biomolecules, such as proteins, nucleic acids, and small molecules. The acrylamide is then polymerized into a hydrogel by a thermally initiated reaction, producing a hybrid construct between tissue and a hydrogel meshwork. An electrophoretic extraction technique then extracts the lipids and replaces them with detergent micelles, leaving the cross-linked molecules in place. The resulting tissue is highly transparent and permits imaging of long-range projections, local wiring, cellular relationships, and subcellular structures in unsectioned tissue. Immunohistochemistry and in situ hybridization are also possible and can be repeated in multiple rounds of differential staining. The technique is also tested successfully on segments of human brain tissue obtained postmortem from a person with autism and is shown to work with brain tissue fixed with formalin. This new technique is anticipated to have a profound effect on the field of brain imaging.

Stadler C, Rexhepaj E, Singan V R, Murphy R F, Pepperkok R, Uhlen M, Simpson J C, Lundberg E. Immunofluorescence and fluorescent-protein tagging show high correlation for protein localization in mammalian cells. *Nature Methods* 10;2013:315–323.

The subcellular localization of proteins may be determined by expressing the protein tagged with a fluorescent marker protein in living cells or by staining fixed cells with fluorescently labeled antibodies against the protein of interest. The former technique is prone to artifacts as a result of the altered structure or overexpression of the protein. The latter technique is susceptible to off-target staining by antibodies of imperfect specificity. The present paper systematically compares localization data acquired by the two techniques for 506 human proteins, the most extensive such study to date. Localization artifacts are observed in the two methods with similar frequency, highlighting the need for cross-validation between the two techniques. The largest discrepancies are observed in dynamic structures, such as endomembranes and cytoskeleton. Nevertheless, 80% of the proteins tested gave concordant results by the two techniques. Nonoverlapping results are most commonly associated with false nuclear localization in fluorescent tagging or immunofluorescence or with incorrect localization to the ER as a result of ectopic overexpression of a fusion protein. The results of the study indicate that the two methods are complementary and should be deployed in tandem.

Yang J, Caprioli R M. Matrix precoated targets for direct lipid analysis and imaging of tissue. *Analytical Chemistry* 85;2013:2907–2912.

The preparation of tissue sections for imaging by MALDI-TOF mass spectrometry has conventionally been performed by applying MALDI matrix to a tissue section immobilized on a metal target plate. There are various methods for matrix application, but none combines convenience, rapidity, and low cost. In the present paper, tissue sections are placed instead on a target precoated with MALDI matrix by sublimation. This procedure is shown to provide images of equivalent quality. The matrix of choice for this procedure is 1,5-diaminonaphthalene, a strongly hydrophobic compound that permits slides to be washed in an aqueous solvent after placement of the tissue section without disrupting the matrix layer. Washing is performed with 50 mM ammonium formate and then with water. Washing removes water-soluble substances, leaving the lipids for imaging. The thicknesses of tissue section and matrix layer are optimized for the procedure. The technique is expected to facilitate mass spectral imaging of tissue lipids.

BIOINFORMATICS

Dutkowski J, Kramer M, Surma M A, Balakrishnan R, Cherry J M, Krogan N J, Ideker T. A gene ontology inferred from molecular networks. *Nature Biotechnology* 31; 2013:38–45.

The Gene Ontology (GO) program has been highly successful in organizing biological knowledge, in terms of the biological processes (e.g., protein synthesis), cellular compartments (e.g., ribosomes), and molecular functions (e.g., elongation factor), of the proteins that genes encode. GO presently contains 35,000 terms and 65,000 relationships between terms to annotate genes from >80 species. However, GO is curated by hand and is therefore susceptible to bias and error. Dutkowski et al. demonstrate that an ontology analogous to GO may be inferred for *Saccharomyces cerevisiae* entirely by automatic computation based on networks derived from high-throughput interaction data: physical protein–protein interactions; genetic interactions, such as synthetic lethality and epistasis; coexpression of genes; and an integrated functional network called YeastNet. To construct an ontology from these networks, the authors build a dendrogram, in which sets of genes with similar interaction patterns are hierarchically joined to-

gether. The gene sets formed in this way are represented by nodes in a tree. These nodes correspond with terms in an ontology. They may be connected by multiway joins rather than just binary joins to reflect the complex relationships assumed by groups of proteins in biology. The network-extracted ontology produced here contains ~4100 biological terms (nodes) and ~5800 biological relationships (edges). The authors compare their computed ontology with GO using an alignment algorithm. They recognize the presence of some entities in the network-extracted ontology have no common English language name, as these entities were not previously recognized by the curators of GO. Such entities can now be incorporated into GO to extend our knowledge of gene function.